

A probabilistic max-plus numerical method for solving stochastic control problems

Marianne Akian¹ and Eric Fodjo²

Abstract—We consider fully nonlinear Hamilton-Jacobi-Bellman equations associated to diffusion control problems involving a finite set-valued (or switching) control and possibly a continuum-valued control. We construct a lower complexity probabilistic numerical algorithm by combining the idempotent expansion properties obtained by McEneaney, Kaise and Han (2011) for solving such problems with a numerical probabilistic method such as the one proposed by Fahim, Touzi and Warin (2011) for solving some fully nonlinear parabolic partial differential equations. Numerical tests on a small example of pricing and hedging an option are presented.

I. INTRODUCTION

We consider a finite horizon diffusion control problem on \mathbb{R}^d involving at the same time a “discrete” control taking its values in a finite set \mathcal{M} , and a “continuum” control taking its values in some subset \mathcal{U} of a finite dimensional space \mathbb{R}^p (for instance a convex set with nonempty interior), which we next describe.

Let T be the horizon. The state $\xi_s \in \mathbb{R}^d$ at time $s \in [0, T]$ satisfies the stochastic differential equation

$$d\xi_s = f^{\mu_s}(\xi_s, u_s)ds + \sigma^{\mu_s}(\xi_s, u_s)dW_s, \quad (1)$$

where $(W_s)_{s \geq 0}$ is a d -dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{0 \leq s \leq T}, P)$. The control processes $\mu := (\mu_s)_{0 \leq s \leq T}$ and $u := (u_s)_{0 \leq s \leq T}$ take their values in the sets \mathcal{M} and \mathcal{U} respectively and they are admissible if they are progressively measurable with respect to the filtration $(\mathcal{F}_s)_{0 \leq s \leq T}$. We assume that, for all $m \in \mathcal{M}$, the maps $f^m : \mathbb{R}^d \times \mathcal{U} \rightarrow \mathbb{R}^d$ and $\sigma^m : \mathbb{R}^d \times \mathcal{U} \rightarrow \mathbb{R}^{d \times d}$ are continuous and satisfy properties implying the existence of the process $(\xi_s)_{0 \leq s \leq T}$ for any admissible control processes μ and u .

Given an initial time $t \in [0, T]$, the control problem consists in maximizing the following payoff:

$$J(t, x, \mu, u) := \mathbb{E} \left[\int_t^T e^{-\int_t^s \delta^{\mu_\tau}(\xi_\tau, u_\tau) d\tau} \ell^{\mu_s}(\xi_s, u_s) ds + e^{-\int_t^T \delta^{\mu_\tau}(\xi_\tau, u_\tau) d\tau} \psi(\xi_T) \mid \xi_t = x \right],$$

where, for all $m \in \mathcal{M}$, $\ell^m : \mathbb{R}^d \times \mathcal{U} \rightarrow \mathbb{R}$, $\delta^m : \mathbb{R}^d \times \mathcal{U} \rightarrow \mathbb{R}_+$ (the set of positive reals), and $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$ are given

continuous maps. We then define the value function of the problem as the optimal payoff:

$$v(t, x) = \sup_{\mu, u} J(t, x, \mu, u),$$

where the maximization holds over all admissible control processes μ and u .

Let \mathbb{S}_d denotes the set of symmetric $d \times d$ matrices. The Hamiltonian $\mathcal{H} : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d \rightarrow \mathbb{R}$ of the above control problem is defined as:

$$\mathcal{H}(x, r, p, \Gamma) := \max_{m \in \mathcal{M}} \mathcal{H}^m(x, r, p, \Gamma),$$

with

$$\mathcal{H}^m(x, r, p, \Gamma) := \max_{u \in \mathcal{U}} \left\{ \frac{1}{2} \text{tr}(\sigma^m(x, u) \sigma^m(x, u)^\top \Gamma) + f^m(x, u) \cdot p - \delta^m(x, u)r + \ell^m(x, u) \right\}.$$

Under suitable assumptions, the value function $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is the unique (continuous) viscosity solution of the following Hamilton-Jacobi-Bellman equation

$$-\frac{\partial v}{\partial t} - \mathcal{H}(x, v(t, x), Dv(t, x), D^2v(t, x)) = 0, \quad (2)$$

$$x \in \mathbb{R}^d, t \in [0, T],$$

$$v(T, x) = \psi(x), \quad x \in \mathbb{R}^d,$$

satisfying also some growth condition at infinity (in space).

In [1], Fahim, Touzi and Warin proposed a probabilistic numerical method to solve such fully nonlinear partial differential equations (2), inspired by their backward stochastic differential equation interpretation given by Cheridito, Soner, Touzi and Victoir in [2]. However this method only works when the diffusion matrices $\sigma^m(x, u) \sigma^m(x, u)^\top$ are at the same time bounded from below (with respect to the Loewner order) by a symmetric positive definite matrix a and bounded from above by $(1 + 2/d)a$. Such a constraint can be restrictive, in particular it may not hold even when the matrices $\sigma^m(x, u)$ do not depend on x and u but take different values for $m \in \mathcal{M}$. Also some regularity conditions may be needed for \mathcal{H} , which are not fulfilled when \mathcal{M} is a finite set.

McEneaney, Kaise and Han proposed in [3], [4] an idempotent numerical method which works at least when the hamiltonian \mathcal{H}^m corresponds to linear quadratic control problems. This method is based on the distributivity of the (usual) addition operation over the supremum (or infimum) operation, and on a property of invariance of the set of quadratic forms. It computes in a backward manner the value function $v(t, \cdot)$ at time t as a supremum of quadratic forms.

¹Marianne Akian is with INRIA and CMAP, École polytechnique CNRS. Address: CMAP, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France. marianne.akian@inria.fr

²Eric Fodjo is with I-Fihn Consulting and INRIA and CMAP, École polytechnique CNRS. Address: CMAP, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France. eric.fodjo@polytechnique.edu

However, as t decreases, the number of quadratic forms generated by the method increases exponentially (and even become infinite if the Brownian is not discretized in space) and some pruning is necessary to reduce the complexity of the algorithm.

Here, we combine the two above methods to construct a new algorithm. The method is using in particular the simulation of a small number of uncontrolled stochastic processes as in [1]. We show that even without pruning, the complexity of the algorithm is bounded polynomially in the number of discretization time steps and in the size of the sample of the uncontrolled stochastic processes. Numerical tests of our algorithm on an example of pricing and hedging an option in dimension 2 considered in [5] are presented.

II. THE ALGORITHM OF FAHIM, TOUZI AND WARIN

Let h be a time discretization step such that T/h is an integer. We denote by $\mathcal{T}_h = \{0, h, 2h, \dots, T-h\}$ the set of discretization times of $[0, T)$.

For each $m \in \mathcal{M}$, we shall assume that we can apply the algorithm of [1] to the equation:

$$-\frac{\partial v}{\partial t} - \mathcal{H}^m(x, v(t, x), Dv(t, x), D^2v(t, x)) = 0, \quad (3)$$

$$x \in \mathbb{R}^d, t \in [0, T).$$

For this purpose, we decompose \mathcal{H}^m as the sum of the (linear) generator \mathcal{L}^m of a diffusion (with no control) and of a nonlinear elliptic Hamiltonian \mathcal{G}^m , that is $\mathcal{H}^m = \mathcal{L}^m + \mathcal{G}^m$ with

$$\mathcal{L}^m(x, r, p, \Gamma) := \frac{1}{2} \text{tr}(a^m(x)\Gamma) + \underline{f}^m(x) \cdot p,$$

$a^m(x) = \sigma^m(x)\sigma^m(x)^\top$ and \mathcal{G}^m such that $\partial_\Gamma \mathcal{G}^m$ is positive semidefinite, for all $x \in \mathbb{R}^d, r \in \mathbb{R}, p \in \mathbb{R}^d, \Gamma \in \mathbb{S}_d$. We also assume that $\text{tr}(a^m(x)^{-1}\partial_\Gamma \mathcal{G}^m) \leq 1$.

The time discretization of (3) proposed in [1] can be written in the following form:

$$v^h(t, x) = T_{t,h}^m(v^h(t+h, \cdot))(x), \quad t \in \mathcal{T}_h,$$

where, under some conditions, $T_{t,h}^m$ is a monotone operator over the set of Lipschitz continuous functions from \mathbb{R}^d to \mathbb{R} :

$$\phi \leq \psi \implies T_{t,h}^m(\phi) \leq T_{t,h}^m(\psi). \quad (4)$$

Moreover, the operator $T_{t,h}^m$ is constructed by using a probabilistic scheme. Denote by \hat{X}^m the Euler discretization of the diffusion with generator \mathcal{L}^m :

$$\hat{X}^m(t+h) = \hat{X}^m(t) + \underline{f}^m(\hat{X}^m(t))h + \underline{\sigma}^m(\hat{X}^m(t))(W_{t+h} - W_t). \quad (5)$$

Then,

$$T_{t,h}^m(\phi)(x) = \mathcal{D}_{m,t,h}^0(\phi)(x) + h\mathcal{G}^m(x, \mathcal{D}_{m,t,h}^0(\phi)(x), \mathcal{D}_{m,t,h}^1(\phi)(x), \mathcal{D}_{m,t,h}^2(\phi)(x)) \quad (6)$$

with, for $i = 0, 1, 2$, $\mathcal{D}_{m,t,h}^i(\phi)$ being the approximation of the i th derivative of ϕ obtained as follows:

$$\mathcal{D}_{m,t,h}^i(\phi)(x) = \mathbb{E}(\phi(\hat{X}^m(t+h)) \mathcal{P}_{m,t,x,h}^i(W_{t+h} - W_t) | \hat{X}^m(t) = x),$$

where, for all m, t, x, h, i , $\mathcal{P}_{m,t,x,h}^i$ is a polynomial of degree i with values in an appropriate finite dimensional space, and in particular $\mathcal{P}_{m,t,x,h}^0 \equiv 1$. Although the operator $T_{t,h}^m$ does not depend on t , since both the law of $W_{t+h} - W_t$ and the Hamiltonian \mathcal{H}^m do not depend on t , we keep the index t since it will become important when applying a regression approximation (see below).

In [1], the convergence of such a time discretization is proved under the above assumptions and some other technical assumptions. Note that these conditions include the boundedness of the coefficients of the Hamiltonians \mathcal{L}^m and \mathcal{H}^m , and the boundedness of the value function of the corresponding control problem. However, a change of variable on the value and on the state allows one to obtain the same type of result for unbounded coefficients and value function satisfying some suitable growth conditions at infinity. Such a change of variables may induce a deformation on the discretizations (5) and (6) and so on the algorithm, but we shall not discuss this here. A greater difficulty is that the above assumptions do not allow in general to handle directly the case where \mathcal{H}^m is replaced by \mathcal{H} . However, in the case of \mathcal{H} as above one can simply consider the following scheme:

$$v^h(t, x) = \max_{m \in \mathcal{M}} \{T_{t,h}^m(v^h(t+h, \cdot))(x)\}, \quad t \in \mathcal{T}_h. \quad (7)$$

The difference with the usual scheme of [1] is that one needs to construct several operators $T_{t,h}^m$ and so several processes \hat{X}^m , one for each $m \in \mathcal{M}$. The solution v^h of this time discretization will converge to the value function of our problem, that is the solution of the Hamilton-Jacobi-Bellman equation with hamiltonian \mathcal{H} , as soon as the convergence is proved for the time discretization of the equations with hamiltonians \mathcal{H}^m .

Although the above scheme can be compared to a standard numerical approximation if one develops the expression of each $T_{t,h}^m(\phi)(x)$, with $m \in \mathcal{M}$, one may compute v^h given by (7) as in [1], that is using a regression estimator. One just simulates the process \hat{X}^m and do at each time $t \in \mathcal{T}_h$ a regression estimation to find the value of $\mathcal{D}_{m,t,h}^i(v^h(t+h, \cdot))$ at the points $\hat{X}^m(t)$ by using the values of $\hat{X}^m(t+h)$ and $W_{t+h} - W_t$.

Although this variation of the method of [1], based on (7), is appealing and may work in practice, several difficulties remain. First, theoretically, the sample size to obtain the convergence of the estimator is at least in the order of $1/h^{d/2}$ [6]. Hence, it is exponential in the dimension of the system showing the persistence of the curse of dimensionality, although in some practical examples, a much smaller sample size may be sufficient. Next, one possible regression estimation is to approximate the conditional expectation of a random map by projecting it orthogonally into a finite dimensional linear space of functions. Then, to obtain a good estimation, the dimension of this space need to be exponential in the dimension d . In the sequel, we shall rather use a small dimensional regression space and use a distributivity property as in the work of McEneaney, Kaise

and Han [3], [4] to find a good approximation of v_h living in the max-plus linear space of finite suprema of quadratic forms.

III. THE ALGORITHM OF MCENEANEY, KAISE AND HAN

In [3], [4], the following time discretization is used. Denote by $\hat{\xi}^{m,u}$ the Euler discretization of the process ξ defined in the introduction, when the controls m and u are fixed:

$$\begin{aligned} \hat{\xi}^{m,u}(t+h) = & \hat{\xi}^{m,u}(t) + f^m(\hat{\xi}^{m,u}(t), u)h + \\ & \sigma^m(\hat{\xi}^{m,u}(t), u)(W_{t+h} - W_t) . \end{aligned}$$

Then, a time discretization of the solution of (2) is given by:

$$v^h(t, x) = T_{t,h}(v^h(t+h, \cdot))(x), \quad t \in \mathcal{T}_h, \quad (8)$$

where

$$\begin{aligned} T_{t,h}(\phi)(x) = & \sup_{m \in \mathcal{M}, u \in \mathcal{U}} \left\{ h\ell^m(x, u) + e^{-h\delta^m(x, u)} \right. \\ & \left. \mathbb{E}[\phi(\hat{\xi}^{m,u}(t+h)) \mid \hat{\xi}^{m,u}(t) = x] \right\}, \end{aligned} \quad (9)$$

Under appropriate assumptions, this scheme converges to the solution of (2) (see [4] for $\delta^m = 0$). Note that the processes \hat{X}^m of the previous section are not related to the processes $\hat{\xi}^{m,u}$ and so the above discretization is different from the one of the previous section.

Assume that the final reward ψ of the control problem can be written as the supremum of a finite number of concave quadratic forms. Denote $\mathcal{Q}_d = \mathbb{S}_d^- \times \mathbb{R}^d \times \mathbb{R}$, where \mathbb{S}_d^- is the set of negative definite symmetric $d \times d$ matrices, and let

$$q(x, z) := \frac{1}{2}x^\top Qx + b \cdot x + c, \quad \text{with } z = (Q, b, c) \in \mathcal{Q}_d, \quad (10)$$

be the quadratic form with parameter z applied to the vector $x \in \mathbb{R}^d$. Then for $g_T = q$, we have

$$v^h(T, x) = \psi(x) = \sup_{z \in Z_T} g_T(x, z)$$

where Z_T is a finite subset of \mathcal{Q}_d . Then, in [4], the following property is deduced from a max-plus distributivity property, in the more general case where the Brownian motion does not have the same dimension as the state space.

Theorem 1 ([4, Theorem 5.1]): Assume that $\delta^m = 0$, that σ^m does not depend on x and u , that f^m is affine with respect to (x, u) , that ℓ^m is concave quadratic with respect to (x, u) , and that ψ is the supremum of a finite number of concave quadratic forms. Consider the time discretization of (8) with (9). Then, for all $t \in \mathcal{T}_h$, there exists a set Z_t and a map $g_t : \mathbb{R}^d \times Z_t \rightarrow \mathbb{R}$ such that for all $z \in Z_t$, $g_t(\cdot, z)$ is a concave quadratic form and

$$v^h(t, x) = \sup_{z \in Z_t} g_t(x, z) .$$

Moreover, the sets Z_t satisfy

$$Z_t = \mathcal{M} \times \{\bar{z}_{t+h} : \mathcal{W} \rightarrow Z_{t+h} \mid \text{Borel measurable}\} ,$$

where $\mathcal{W} = \mathbb{R}^d$ is the space of values of the Brownian process.

Note that the sets Z_t are infinite as soon as $t < T$. However, if the Brownian process is discretized in space, the set \mathcal{W} can be replaced by a finite subset, and the sets Z_t become finite. Nevertheless, their cardinality increases exponentially as t decreases: $\#Z_t = \#\mathcal{M} \times (\#Z_{t+h})^p$ where p is the cardinality of the discretization of \mathcal{W} . Then, McEneaney, Kaise and Han proposed in [4] a pruning method to reduce at each time step $t \in \mathcal{T}_h$ the cardinality of Z_t .

IV. COMBINING MAX-PLUS APPROXIMATIONS AND PROBABILISTIC SCHEMES

Here, we assume that the assumptions of the two previous sections hold, and consider a time discretization scheme similar to the one of Section II. The application of the operator $T_{t,h}^m$ of (6) to a function $\phi : \mathbb{R}^d \rightarrow \mathbb{R}, x \mapsto \phi(x)$ can be written, for each $x \in \mathbb{R}^d$, as

$$T_{t,h}^m(\phi)(x) = G_{t,x,h}^m(\tilde{\phi}_{t,x,h}^m), \quad (11a)$$

where $G_{t,x,h}^m$ is an operator from \mathcal{D} to \mathbb{R} , where \mathcal{D} is the set of measurable functions from $\mathcal{W} := \mathbb{R}^d$ to \mathbb{R} with at most exponential growth rate, and

$$\tilde{\phi}_{t,x,h}^m : \mathcal{W} \rightarrow \mathbb{R}, W \mapsto \phi(x + \underline{f}^m(x)h + \underline{\sigma}^m(x)W) . \quad (11b)$$

Indeed in the case of (6), $G_{t,x,h}^m$ is given by

$$\begin{aligned} G_{t,x,h}^m(\tilde{\phi}) = & D_{m,t,x,h}^0(\tilde{\phi}) \\ & + h\mathcal{G}^m(x, D_{m,t,x,h}^0(\tilde{\phi}), D_{m,t,x,h}^1(\tilde{\phi}), D_{m,t,x,h}^2(\tilde{\phi})) \end{aligned} \quad (12)$$

with, for $i = 0, 1, 2$,

$$D_{m,t,x,h}^i(\tilde{\phi}) = \mathbb{E}(\tilde{\phi}(W_{t+h} - W_t) \mathcal{P}_{m,t,x,h}^i(W_{t+h} - W_t)) .$$

Let us say that an operator $G : \mathcal{D} \rightarrow \mathbb{R}$ is monotone if it satisfies:

$$\phi, \psi \in \mathcal{D}, \phi \leq \psi \text{ a.e.}, \implies G(\phi) \leq G(\psi) , \quad (13a)$$

and that it is additively α -subhomogeneous, for some constant $\alpha > 0$, if it satisfies:

$$\forall K > 0 \text{ and } \phi \in \mathcal{D}, G(\phi + K) \leq G(\phi) + \alpha K , \quad (13b)$$

where $\phi + K$ is the map $W \in \mathcal{W} \mapsto \phi(W) + K$. The above properties imply that the restriction of G to the set of bounded measurable functions is Lipschitz continuous with constant α . Using the same kind of proof as in [1] for (4), one can obtain the stronger property that all the operators $G_{t,x,h}^m$ belong to the class of monotone additively α_h -subhomogeneous operators from \mathcal{D} to \mathbb{R} , for some constant $\alpha_h = 1 + Ch$ with $C \geq 0$. This implies that $T_{t,h}^m$ sends the set of bounded measurable functions to itself and is Lipschitz continuous with constant α_h on it.

Since we shall consider approximations of the value function by suprema of concave quadratic forms, we shall need to apply the operators $T_{t,h}^m$ to quadratic forms, which are unbounded maps. Therefore, in order to apply the above

properties, we shall rather assume that the operators $T_{t,h}^m$ can also be written as

$$T_{t,h}^m(\phi) = \tilde{T}_{t,h}^m(\phi/\chi)\chi \quad (14)$$

where $\chi : \mathbb{R}^d \rightarrow \mathbb{R}$ is given by $\chi(x) = 1 + \|x\|^2$ with $\|\cdot\|$ being the Euclidian norm, and where $\tilde{T}_{t,h}^m$ are also of the above form (11) with some operators $\tilde{G}_{t,x,h}^m$ instead of $G_{t,x,h}^m$ belonging again to the class of monotone additively $\tilde{\alpha}_h$ -subhomogeneous operators from \mathcal{D} to \mathbb{R} , for some constant $\tilde{\alpha}_h = 1 + \tilde{C}h$ with $\tilde{C} \geq 0$. Under these conditions, We obtain a result similar to [4, Theorem 5.1].

Theorem 2: Consider the control problem of Section I. Assume that δ^m and σ^m are constant, that f^m is affine with respect to (x, u) , that ℓ^m is concave quadratic with respect to (x, u) , and that ψ is the supremum of a finite number of concave quadratic forms. Consider the time discretization (7) with $\underline{\sigma}^m = \varepsilon \sigma^m$, $0 < \varepsilon \leq 1$ and \underline{f}^m affine. Assume that the operators $T_{t,h}^m$ satisfy (14) for some operators $\tilde{T}_{t,h}^m$ of the form (11), with some operators $\tilde{G}_{t,x,h}^m$ instead of $G_{t,x,h}^m$ belonging to the class of monotone additively $\tilde{\alpha}_h$ -subhomogeneous operators from \mathcal{D} to \mathbb{R} , for some constant $\tilde{\alpha}_h = 1 + \tilde{C}h$ with $\tilde{C} \geq 0$, see (13). Assume also that the discretized value function v_h of (7) is such that v_h/χ is bounded and Lipschitz continuous with respect to x . Then, for all $t \in \mathcal{T}_h$, there exists a set Z_t and a map $g_t : \mathbb{R}^d \times Z_t \rightarrow \mathbb{R}$ such that for all $z \in Z_t$, $g_t(\cdot, z)$ is a concave quadratic form and

$$v^h(t, x) = \sup_{z \in Z_t} g_t(x, z) .$$

Moreover, the sets Z_t satisfy

$$Z_t = \mathcal{M} \times \{\bar{z}_{t+h} : \mathcal{W} \rightarrow Z_{t+h} \mid \text{Borel measurable}\} .$$

This result uses the following properties, the second one being a generalization of [4, Theorem 3.1].

Lemma 3: Let \bar{z} be a measurable function from \mathcal{W} to \mathcal{D}_d . Let us consider the notations and assumptions of Theorem 2 and let $G_{t,x,h}^m$ be related to $T_{t,h}^m$ by (12). Let $\tilde{q}_{t,x,h}^{m,\bar{z}}$ be the map $\mathcal{W} \rightarrow \mathbb{R}$, $W \mapsto q(x + \underline{f}^m(x)h + \sigma^m(x)W, \bar{z}(W))$, with q as in (10). Then, the function $x \mapsto G_{t,x,h}^m(\tilde{q}_{t,x,h}^{m,\bar{z}})$ is a concave quadratic form, that is it can be written as $q(x, Z)$ for some $Z \in \mathcal{D}_d$.

Theorem 4: Let $\mathcal{W} = \mathbb{R}^d$ and G be a monotone additively α -subhomogeneous operator from \mathcal{D} to \mathbb{R} , for some constant $\alpha > 0$, see (13). Let (Z, \mathfrak{A}) be a measurable space, and let \mathcal{W} be endowed with its Borel σ -algebra. Let $\phi : \mathcal{W} \times Z \rightarrow \mathbb{R}$ be measurable map such that for all $z \in Z$, $\phi(\cdot, z)$ is bounded and continuous. Let $v : \mathcal{W} \rightarrow \mathbb{R}$ be such that $v(W) = \sup_{z \in Z} \phi(W, z)$. Assume that v is continuous and bounded. Then,

$$G(v) = \sup_{\bar{z} \in \bar{Z}} G(\bar{\phi}^{\bar{z}})$$

where $\bar{\phi}^{\bar{z}} : \mathcal{W} \rightarrow \mathbb{R}$, $W \mapsto \phi(W, \bar{z}(W))$, and

$$\bar{Z} = \{\bar{z} : \mathcal{W} \rightarrow Z, \text{ measurable} \\ \text{and such that } \bar{\phi}^{\bar{z}} \text{ is bounded}\} .$$

Proof: Since v is bounded and continuous, it belongs to \mathcal{D} , so that $G(v)$ is well defined. Similarly, by definition,

for all $\bar{z} \in \bar{Z}$, $\bar{\phi}^{\bar{z}}$ is measurable and bounded, so it belongs to \mathcal{D} , so that $G(\bar{\phi}^{\bar{z}})$ is well defined.

Let $\varepsilon > 0$. By definition of v , for all $W \in \mathcal{W}$, there exists $z^W \in Z$ such that $\phi(W, z^W) \geq v(W) - \varepsilon$. Then, since $W' \mapsto \phi(W', z^W)$ and $W' \mapsto v(W')$ are continuous maps $\mathcal{W} \rightarrow \mathbb{R}$, there exists $\delta^W > 0$ such that for all $W' \in B(W, \delta^W)$ (the open ball centered at W with radius δ^W), $|\phi(W', z^W) - \phi(W, z^W)| \leq \varepsilon$ and $|v(W') - v(W)| \leq \varepsilon$. Then, for $W' \in B(W, \delta^W)$, we have

$$\phi(W', z^W) \geq \phi(W, z^W) - \varepsilon \geq v(W) - 2\varepsilon \geq v(W') - 3\varepsilon .$$

As \mathcal{W} is the countable union of compact metric spaces, there exists a sequence $(W_i)_{i \geq 0}$ of \mathcal{W} such that $\mathcal{W} = \bigcup_{i \geq 0} B(W_i, \delta^{W_i})$. Let us denote, for all $i \geq 0$, $\mathcal{W}_i = B(W_i, \delta^{W_i})$ and $\mathcal{W}'_i = \mathcal{W}_i \setminus (\bigcup_{j < i} \mathcal{W}_j)$. Define the function \bar{z} such that, for all $i \geq 0$, $\bar{z}(W') = z^{W_i}$, for $W' \in \mathcal{W}'_i$. Since $(\mathcal{W}'_i)_{i \geq 0}$ is a countable partition of \mathcal{W} composed of Borel sets, the map \bar{z} is well defined on \mathcal{W} and measurable. Moreover, by the above properties and the definition of v , we have

$$v(W) \geq \bar{\phi}^{\bar{z}}(W) = \phi(W, \bar{z}(W)) \geq v(W) - 3\varepsilon, \quad \forall W \in \mathcal{W} .$$

Since v is bounded, this implies that $\bar{\phi}^{\bar{z}}$ is bounded, which implies that \bar{z} belongs to \bar{Z} .

Since G is monotone and additively α -subhomogeneous from \mathcal{D} to \mathbb{R} , and $\varepsilon > 0$, we get that

$$G(v) \geq G(\bar{\phi}^{\bar{z}}) \geq G(v - 3\varepsilon) \geq G(v) - 3\alpha\varepsilon .$$

Then

$$G(v) \geq \sup_{\bar{z} \in \bar{Z}} G(\bar{\phi}^{\bar{z}}) \geq G(v) - 3\alpha\varepsilon ,$$

and since this property holds for all $\varepsilon > 0$, we obtain the equality, which shows the assertion of the theorem. ■

Using Theorem 2, we get that $v^h(t, \cdot)$ is the supremum of concave quadratic maps, but as in [4, Theorem 5.1], the sets Z_t are infinite for $t < T$. Here, we shall compute the expression of the maps $v^h(t, \cdot)$ by approximating the operators $T_{t,h}^m$ as in the same spirit as in [1], that is using the simulation of the processes \hat{X}^m . The main difference with the method of [1] is that the egression estimations are done on quadratic forms and not on the value functions directly. Because of the simulations, we should only need to compute the values $v^h(t, \hat{X}^m(t))$. This means that if N is the number of samples of the Brownian process, then the number of quadratic forms $g_t(\cdot, z)$ that are essentials in the computation of $v^h(t, \hat{X}^m(t))$ as a supremum of quadratic forms is less or equal to $N \times M$, where M is the cardinality of \mathcal{M} . Then, for any random quadratic form which is optimal for a particular $\hat{X}^m(t)$, we need to compute its image by $T_{t,h}^m$. In [4], this image is obtained by hand. Here, we shall rather use a regression estimation. The result of Lemma 3 implies that taking for the linear regression space, the space of quadratic forms gives an exact result at least when the sample size is large (to ensure that the solution of the estimation problem is unique).

With all these properties in mind, we construct the following algorithm. Let us denote by $S_{t,h}^m(x, w)$ the following

operator which sends $(\hat{X}^m(t), W_{t+h} - W_t)$ into $\hat{X}^m(t+h)$:

$$S_{t,h}^m(x, w) = x + \underline{f}^m(x)h + \underline{\sigma}^m(x)w.$$

Since $\underline{\sigma}^m$ is constant and \underline{f}^m is affine, the map $S_{t,h}^m(\cdot, w)$ is affine for all $w \in W$.

Algorithm 1:

Input: A constant ε giving the precision, and a 5-uple $N = (N_{\text{in}}, N_{\text{rg}}, N_x, N_w, N_m)$ of integers giving the numbers of samples and the “method of sampling” $N_m \in \{1, \dots, 5\}$ described below. A finite subset Z_T of \mathcal{Q}_d such that $|\psi(x) - \max_{z \in Z_T} q(x, z)| \leq \varepsilon$, for all $x \in \mathbb{R}^d$, and $\#Z_T \leq M \times N_{\text{in}}$, and the operators $T_{t,h}^m$ and $G_{t,x,h}^m$ as in Theorem 2.

Output: The subsets Z_t of \mathcal{Q}_d , for $t \in \mathcal{T}_h \cup \{T\}$, and the approximate value function $v^{h,N} : (\mathcal{T}_h \cup \{T\}) \times \mathbb{R}^d \rightarrow \mathbb{R}$.

• *Initialization:* Let $\hat{X}^m(0)$ be random and independent of the Brownian process. Consider a sample of $(\hat{X}^m(0), (W_{t+h} - W_t)_{t \in \mathcal{T}_h})$ of size N_{in} indexed by $\omega \in \Omega_{N_{\text{in}}} := \{1, \dots, N_{\text{in}}\}$, and denote, for each $t \in \mathcal{T}_h \cup \{T\}$ and $\omega \in \Omega_{N_{\text{in}}}$, $\hat{X}^m(t, \omega)$ the value of $\hat{X}^m(t)$ induced by this sample and satisfying (5). Define $v^{h,N}(T, x) = \max_{z \in Z_T} q(x, z)$, for $x \in \mathbb{R}^d$, with q as in (10).

• For $t = T - h, T - 2h, \dots, 0$ apply the following 3 steps:

(1) For each $\omega \in \Omega_{N_{\text{in}}}$ and $m \in \mathcal{M}$, construct a sample $(\omega_1, \omega'_1), \dots, (\omega_{N_{\text{rg}}}, \omega'_{N_{\text{rg}}})$ of elements of $\Omega_{N_{\text{in}}} \times \Omega_{N_{\text{in}}}$, using the method N_m and possibly the constants N_x and N_w . Induce the sample $\hat{X}^m(t, \omega_i)$ (resp. $(W_{t+h} - W_t)(\omega'_i)$) for $i \in \Omega_{N_{\text{rg}}}$ of $\hat{X}^m(t)$ (resp. $W_{t+h} - W_t$). Denote by $\mathcal{W}_t^N \subset \mathcal{W}$ the set of $(W_{t+h} - W_t)(\omega'_i)$ for $i \in \Omega_{N_{\text{rg}}}$.

(2) For each $\omega \in \Omega_{N_{\text{in}}}$ and $m \in \mathcal{M}$, construct $z_t \in \mathcal{Q}_d$ depending on ω and m as follows:

Let $\bar{z}_{t+h} : \mathcal{W}_t^N \rightarrow Z_{t+h} \subset \mathcal{Q}_d$ be such that, for all $i \in \Omega_{N_{\text{rg}}}$ we have

$$\begin{aligned} v^{h,N}(t+h, S_{t,h}^m(\hat{X}^m(t, \omega), (W_{t+h} - W_t)(\omega'_i))) \\ = q(S_{t,h}^m(\hat{X}^m(t, \omega), (W_{t+h} - W_t)(\omega'_i)), \bar{z}_{t+h}((W_{t+h} - W_t)(\omega'_i))) \end{aligned}$$

Extend \bar{z}_{t+h} as a measurable map on \mathcal{W} . Let $\tilde{q}_{t,x,h}^{m,\bar{z}}$ be as in Lemma 3, that is be the map $\mathcal{W} \rightarrow \mathbb{R}$, $W \mapsto q(x + \underline{f}^m(x)h + \underline{\sigma}^m(x)W, \bar{z}(W))$. Compute an approximation of $x \mapsto G_{t,x,h}^m(\tilde{q}_{t,x,h}^{m,\bar{z}})$ by a regression estimation on the set of quadratic forms using the sample $(\hat{X}^m(t, \omega_i), (W_{t+h} - W_t)(\omega'_i))$, with $i \in \Omega_{N_{\text{rg}}}$. We obtain $z_t \in \mathcal{Q}_d$ such that $q(x, z_t) \simeq G_{t,x,h}^m(\tilde{q}_{t,x,h}^{m,\bar{z}})$.

(3) Denote by Z_t the set of all the $z_t \in \mathcal{Q}_d$ obtained in this way, and define

$$v^{h,N}(t, x) = \max_{z \in Z_t} q(x, z) \quad \forall x \in \mathbb{R}^d.$$

Let us precise now the different choices of the “method of sampling” N_m used in the algorithm:

Method 1: Assume $N_{\text{rg}} = N_{\text{in}}$ and take $\omega_i = \omega'_i = i$ for $i \in \Omega_{N_{\text{rg}}}$, which means that we take the initial sampling.

Method 2: Assume $N_{\text{rg}} = N_x \times N_w$, and choose once for all $\omega \in \Omega_{N_{\text{in}}}$ and $m \in \mathcal{M}$ in the algorithm: a random sampling $\omega_{i,1}$, $i = 1, \dots, N_x$ among the elements of $\Omega_{N_{\text{in}}}$ and independently a

random sampling $\omega'_{1,j}$, $j = 1, \dots, N_w$ among the elements of $\Omega_{N_{\text{in}}}$, then take the product of samplings, leading to $(\omega_{i,1}, \omega'_{1,j})$ for $i = 1, \dots, N_x$ and $j = 1, \dots, N_w$. Reindexing the sampling, we obtain (ω_i, ω'_i) for $i = 1, \dots, N_{\text{rg}}$.

Method 3: Do as in Method 2, but choose different samplings for each $\omega \in \Omega_{N_{\text{in}}}$ and $m \in \mathcal{M}$ in the algorithm, independently.

Method 4: Assume $N_{\text{rg}} = N_x \times N_w$ and $N_w = N_{\text{in}}$ and do as in Method 2, but take the fixed sampling $\omega'_{1,j} = j$ instead of random sampling.

Method 5: Assume $N_{\text{rg}} = N_{\text{in}}^2$ and do as in Method 2, but take the fixed samplings $\omega_{i,1} = i$ and $\omega'_{1,j} = j$ instead of random samplings.

It is easy to see that the sets Z_t of the above algorithm satisfy $\#Z_t \leq M \times N_{\text{in}}$ for all $t \in \mathcal{T}_h$. Then, the number of computations at each time step for the optimization (computation of the \bar{z}_{t+h}) will be at most in the order of $(M \times N_{\text{in}})^2 \times N_{\text{rg}}$ and at most in the order of $(M \times N_{\text{in}})^2 \times N_w$ when using methods 2,3,4. Moreover, the number of computations at each time step for the regression estimation will be at most in the order of $\mathcal{M} \times N_{\text{in}} \times N_{\text{rg}}$ so will be negligible with respect to the optimization step.

Note that in the above algorithm, the regression estimation depends only on the value of \bar{z}_{t+h} on the simulations $(W_{t+h} - W_t)(\omega_i)$, with $i \in \Omega_{N_{\text{rg}}}$. That is the extension of \bar{z}_{t+h} to a measurable function on \mathcal{W} was only needed for the definition of $\tilde{q}_{t,x,h}^{m,\bar{z}}$ and $x \mapsto G_{t,x,h}^m(\tilde{q}_{t,x,h}^{m,\bar{z}})$. From Lemma 3, $p : x \mapsto G_{t,x,h}^m(\tilde{q}_{t,x,h}^{m,\bar{z}})$ is a quadratic form. Hence, when N_{rg} is large, the regression estimation of p as a quadratic form using the given sample is a good approximation. Under these conditions, we have $v^{h,N}(t, x) = \sup_{m \in \mathcal{M}} v^{h,N,m}(t, x)$ for all x , with $v^{h,N,m}(t, x) \simeq T_{t,h}^m(v^{h,N}(t+h, \cdot))(x)$ holding for all $x = \hat{X}^m(t, \omega)$ with $\omega \in \Omega_{N_{\text{in}}}$. This implies that, when the regression approximations converge, $v^{h,N}$ is a good approximation of v^h . Then, under the assumptions of Theorem 2, one may expect a convergence result comparable to the one of [1], showing the existence of some 5-uples N_h such that v^{h,N_h} converges towards the value function of the control problem when h goes to 0. We present numerical tests to confirm this convergence in the next section, at least in the case of some of the sampling methods proposed above. The precise convergence study is left for further work.

V. NUMERICAL TESTS

To test our algorithm, we consider the problem of pricing and hedging an option with uncertain volatility and two underlying processes, studied as an example in Section 3.2 of [5]. There, the method proposed is based on a regression on a process involving not only the state but also the (discrete) control.

With the notations of the introduction, we consider the case where $d = 2$, $\mathcal{M} = \{\rho_{\min}, \rho_{\max}\}$ with $-1 \leq \rho_{\min}, \rho_{\max} \leq 1$, and there is no continuum control, so u is omitted. The dynamics of the processes are given, for all $m \in \mathcal{M}$, by

$f^m = 0$, and for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$,

$$\sigma^m(\xi) = \begin{bmatrix} \sigma_1 \xi_1 & 0 \\ \sigma_2 m \xi_2 & \sigma_2 \sqrt{1 - m^2} \xi_2 \end{bmatrix}$$

with $\sigma_1, \sigma_2 > 0$. The parameters of the reward satisfy $\delta^m = 0$, $l^m = 0$, and, for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$,

$$\psi(\xi) = (\xi_1 - \xi_2 - K_1)^+ - (\xi_1 - \xi_2 - K_2)^+$$

with $x^+ = \max(x, 0)$, $K_1 < K_2$.

The two coordinates of the controlled process stay in \mathbb{R}_+ , the set of positive reals. To be in the conditions of Theorem 2, we approximate the function ψ with a supremum of a finite number of concave quadratic forms on a large subset of \mathbb{R}_+^2 , typically on the set of ξ such that $\xi_1 - \xi_2 \in [-100, 100]$. Note that since the second derivative of ψ is $-\infty$ in some points, it is not c -semiconvex for any $c > 0$ and bounded domain, so the approximation need to use some quadratic forms with a large negative curvature, and so the algorithm proposed in [4] may not work. The maps σ^m for $m \in \mathcal{M}$ are not constant but they are linear, so one can show that the result of Theorem 2 still holds.

We take the same constants as in [5]: $\sigma_1 = 0.4$, $\sigma_2 = 0.3$, $K_1 = -5$, $K_2 = 5$, $T = 0.25$, $\rho_{\min} = -0.8$, $\rho_{\max} = 0.8$. We fix the time discretization step to $h = 0.1$.

We first tested our algorithm in the case where \mathcal{M} is the singleton $\{\rho_{\min}\}$ or $\{\rho_{\max}\}$, which means that there is no action on the process, so that the true value function can be computed analytically, and compared with the solution obtained by our algorithm. The method $N_m = 1$ gives very bad results even at time $T - h$. The method $N_m = 5$ need too much space and time even for $N_{\text{in}} = 1000$. In Table V, we present for different values of $N = (N_{\text{in}}, N_{\text{rg}}, N_x, N_w, N_m)$, with $N_m = 2, 3, 4$, the norm of the error on the value function at time $t = 0$ and states $\xi_2 = 50$ and $\xi_1 \in [20, 80]$. We see that the best method is the second one, and that Method 3 gives very bad results. This may be explained by the introduction of a biais due to the maximization of independent random variables. Note also that the errors for Method 2 are comparable to the standard deviations obtained in [7] by Gobet, Lemor and Warin in the case of similar option problems with a usual regression estimation of the value function.

In view of these results, we present in Figure 1 the result obtained for the control problem tested in [5], that is with $\mathcal{M} = \{\rho_{\min}, \rho_{\max}\}$, and $N_{\text{in}} = 1000$, $N_{\text{rg}} = N_x \times N_w$, $N_x = 10$, $N_w = 1000$ and $N_m = 2$. The result is very similar to the one presented in [5].

REFERENCES

- [1] A. Fahim, N. Touzi, and X. Warin, "A probabilistic numerical method for fully nonlinear parabolic PDEs," *Ann. Appl. Probab.*, vol. 21, no. 4, pp. 1322–1364, 2011. [Online]. Available: <http://dx.doi.org/10.1214/10-AAP723>
- [2] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir, "Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs," *Comm. Pure Appl. Math.*, vol. 60, no. 7, pp. 1081–1110, 2007. [Online]. Available: <http://dx.doi.org/10.1002/cpa.20168>

ρ	N_{in}	N_{rg}	N_x	N_w	N_m	e_∞	e_1
-0.8	1000	10000	10	1000	2	0.521	0.173
0.8	1000	10000	10	1000	2	0.157	0.074
-0.8	1000	1000	10	100	2	0.75	0.41
0.8	1000	1000	10	100	2	0.36	0.11
-0.8	1000	1000	10	100	3	3.48	1.92
0.8	1000	1000	10	100	3	3.05	0.81
-0.8	100	1000	10	100	2	1.95	0.46
0.8	100	1000	10	100	2	1.81	0.33
-0.8	100	10000	10	1000	2	2.09	0.53
0.8	100	10000	10	1000	2	1.79	0.36
-0.8	100	1000	10	100	4	2.15	0.55
0.8	100	1000	10	100	4	1.80	0.39

TABLE I

SUP-NORM AND NORMALIZED ℓ^1 NORM OF THE ERROR, ON THE VALUE FUNCTION WITH CONSTANT ρ , AT TIME $t = 0$, AND STATES $\xi_2 = 50$ AND $\xi_1 \in [20, 80]$, DENOTED e_∞ AND e_1 RESP.

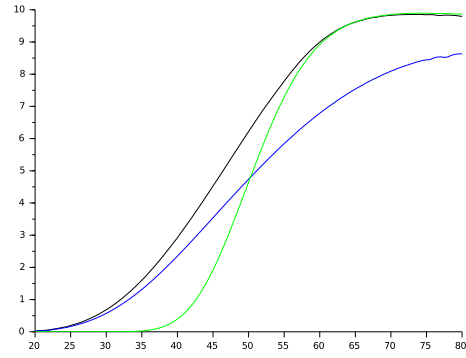


Fig. 1. Value function obtained at $t = 0$, and $\xi_2 = 50$ as a function of $\xi_1 \in [20, 80]$. Here $N_{\text{in}} = 1000$, $N_{\text{rg}} = N_x \times N_w$, $N_x = 10$, $N_w = 1000$ and $N_m = 2$. In blue, ρ is constant equal to -0.8 , in green ρ is constant equal to 0.8 , and in black $\rho \in \{-0.8, 0.8\}$.

- [3] H. Kaise and W. M. McEneaney, "Idempotent expansions for continuous-time stochastic control: compact control space," in *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, Dec. 2010.
- [4] W. M. McEneaney, H. Kaise, and S. H. Han, "Idempotent method for continuous-time stochastic control and complexity attenuation," in *Proceedings of the 18th IFAC World Congress, 2011*, Milano, Italie, 2011, pp. 3216–3221.
- [5] I. Kharroubi, N. Langrené, and H. Pham, "A numerical algorithm for fully nonlinear HJB equations: an approach by control randomization," *Monte Carlo Methods Appl.*, vol. 20, no. 2, pp. 145–165, 2014. [Online]. Available: <http://dx.doi.org/10.1515/mcma-2013-0024>
- [6] B. Bouchard and N. Touzi, "Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations," *Stochastic Process. Appl.*, vol. 111, no. 2, pp. 175–206, 2004. [Online]. Available: <http://dx.doi.org/10.1016/j.spa.2004.01.001>
- [7] E. Gobet, J.-P. Lemor, and X. Warin, "A regression-based monte carlo method to solve backward stochastic differential equations," *The Annals of Applied Probability*, vol. 15, no. 3, p. 21722202, 2005.